The Interviewer in the Respondent's Shoes: What Can We Learn from the Way Interviewers Answer Survey Questions?

Celine Wuyts*

Centre for Sociological Research, KU Leuven

celine.wuyts@soc.kuleuven.be

Geert Loosveldt

Centre for Sociological Research, KU Leuven

geert.loosveldt@soc.kuleuven.be

*Corresponding Author:
Celine Wuyts, Centre for Sociological Research, KU Leuven,
Parkstraat 45, 3000 Leuven, Belgium.
Email: celine.wuyts@soc.kuleuven.be

Research funded by Hercules Foundation Flanders

Abstract

Previous research shows that interviewers to some extent fail to expend the effort that is needed to collect high-quality survey data. We extend the idea of interviewer satisficing to a related task, in which the interviewers themselves answer survey questions. We hypothesize that interviewers who self-administer the questionnaire in a careless manner, also will not apply themselves fully to the task of administering survey interviews. Based on interviewer and respondent data from the sixth round of the European Social Survey in Belgium, we find support for some of the hypothesized associations between (suboptimal) response characteristics of interviewers in the "task as respondent" and the same (suboptimal) response characteristics recorded for their respondents, specifically with regard to interview speed, multiple response, and item nonresponse to the household income question.

Keywords: Interviewer satisficing, interviewer behavior, European Social Survey

Introduction

Even though the quality of survey data depends on it, neither interviewers nor respondents necessarily perform their tasks optimally. Rather than being optimal, much human behavior is based on satisfactory or acceptable choices (Simon 1957). This tendency to *satisfice* has been used to explain the behavior of respondents (Krosnick and Alwin 1987; Krosnick, Narayan, and Smits 1996; Krosnick 2000) and interviewers (Japec 2008; Ongena and Dijkstra 2007) during survey interviews. A survey respondent can produce an optimal answer by carrying out a series of cognitive steps thoroughly (Tourangeau, Rips, and Rasinski 2000). If the respondent is unable or unwilling to expend the cognitive effort required to provide an optimal response, he or she may compromise or skip one or more of these steps, and provide a minimally acceptable response instead (Krosnick and Alwin 1987; Krosnick, Narayan, and Smits 1996). Analogously, asking questions, probing, and recording answers also requires the interviewer to thoroughly execute a series of cognitive steps, which he or she may fail to do (Ongena and Dijkstra 2007). Studies have shown that interviewers might partly deviate from question wording and interviewing instructions (e.g. Cannell, Fowler, and Marquis 1968; Mangione, Fowler, and Louis 1992; Dijkstra and Ongena 2006; Ackermann-Piek and Massing 2014). Differences in interviewers' task performance is also evident from observed interviewer effects on survey estimates (O'Muircheartaigh and Campanelli 1998; West, Kreuter, and Jaenichen 2013), item nonresponse (Pickery and Loosveldt 1998, 2001), and interview speed (Loosveldt and Beullens 2013). Although interviewers may feel compelled to deviate in order to make the conversation flow more naturally (Houtkoop-Steenstra 2005) and collect more accurate answers (Schober and Conrad

1997), they may also do so to reduce their own burden (Ongena and Dijkstra 2007; Japec 2008). If interviewers satisfice during their tasks, the characteristics of respondents' answers that are usually attributed to suboptimal response behavior–such as item nonresponse and non-differentiation–might to some extent be due to suboptimal behavior on the part of the interviewers.

The concept of satisficing can also be extended to an interviewer task different from, but related to, the tasks in an actual survey interview (in which the interviewer interacts *with* respondents), namely the task of self-administering a questionnaire (in which the interviewer acts *as* the respondent). By completing the respondent questionnaire, an interviewer familiarizes him or herself with the survey instrument prior to the start of the fieldwork (Billiet and Loosveldt 1988), instead of during the first few interviews, as suggested by an initial increase in interview speed (Olson and Peytchev 2007). The "task as respondent" is thus valuable in and of itself. Moreover, the collected data allows researchers and practitioners to study the interviewers' own response behavior in an indirect way. The general research questions is what can we learn from the way interviewers answer survey questions. The task puts the interviewer "in the respondent's shoes," and similarly to any other respondent, interviewers can satisfice in order to minimize their burden. Hence, we can distinguish two types of interviewer satisficing: the interviewer satisficing during an actual interview, resulting in straying from instructions, and interviewer satisficing in the "task as respondent," resulting in suboptimal response behavior when they complete the questionnaire.

Assuming that people exhibit a general maximizing versus satisficing orientation (Schwartz 2002), satisficing behavior will be expressed in different tasks,

instead of being purely circumstantially determined. We therefore propose that interviewers who expend little effort when answering survey questions themselves, will also expend little effort when interacting with respondents in the actual interview process, as well as tolerating or encouraging respondents' satisficing behavior. We examine this proposition by considering a selected number of formal response characteristics that have been related to respondent satisficing behavior–i.e. item nonresponse, multiple response, differentiation, and interview speed–for interviewers who self-administered the questionnaire as well as for the actual survey respondents.

Respondent item nonresponse, multiple response, differentiation, and interview speed are hypothesized to be positively associated with respectively interviewer item nonresponse, interviewer multiple response, interviewer differentiation, and interviewer interview speed. With the exception of item nonresponse, these formal characteristics of interviewers' own responses have not, to our knowledge, previously been included in the analysis of interviewer effects on data quality. The effects of interviewers' item nonresponse were studied by Pickery and Loosveldt (1998, 2001), with mixed results. No statistically significant effect of the interviewers' item nonresponse (as measured by the number of "no opinion" answers) on the respondents' item nonresponse was found in the Belgian 1991 post-election survey (Pickery and Loosveldt 1998). However, in the 1995 post-election survey, Pickery and Loosveldt (2001) found that for several questions on politics, item nonresponse by respondents was more likely if the interviewer had not answered the respective question themselves. This finding suggests that interviewers who make the effort to complete all questions themselves, consistently probe "don't know" or "no opinion" answers to a greater extent. Similar mechanisms may link interviewers'

multiple response, differentiation, and interview speed to their respondents' multiple response, differentiation, and interview speed, respectively. Interviewers who consider all the response options to "all that apply" questions are more likely to select more than one answer in the self-administered questionnaire (Rasinski, Mingay, and Bradburn 1994) and may consistently probe for additional answers, resulting in more answers given on average. Interviewers who consider different items in a grid separately, are less likely to straight line and are expected to put more effort into motivating respondents to do the same. Lastly, interviewers who take more time to complete the questionnaire themselves will also read questions more slowly and/or allow respondents more time to answer, resulting in lower interview speed.

If the way in which the interviewers complete the "task as respondent" affects data quality, survey practitioners could benefit from evaluating this task and adapting fieldwork preparation accordingly. Project briefings might be designed to pay more attention to the instructions from which the interviewers are likely to deviate.

Data

The current study is based on data from round six of the European Social Survey (ESS) (European Social Survey 2012), fielded in Belgium in the fall of 2012, and interviewer data from the round six questionnaire. The questionnaire was administered to survey respondents using computer-assisted personal interviewing. Following the principles of standardized interviewing, the interviewers were supposed to read the questions exactly as worded, probe in a neutral way when necessary, and enter the answers exactly as given on their laptop or tablet. Before attending the ESS project briefing, the interviewers were asked to complete and submit the main

questionnaire in order to get a feel for the instrument. This "task as respondent" resulted in 155 self-administered questionnaires from the interviewers, in addition to the 1,869 questionnaires completed by actual respondents.

Methods

Measurements

Based on existing literature and the specifics of the ESS6 questionnaire, we selected eight formal characteristics of response, covering item nonresponse, multiple response, non-differentiation, and interview speed. These response indicators were calculated for the respondents being personally interviewed and for the interviewers who self-administered the questionnaire (the "task as respondent") as described in the following paragraphs.

In line with Pickery and Loosveldt (2001), binary item nonresponse indicators were constructed. Item nonresponse in the Belgian ESS6 data is generally low, and concentrated in just a few questions, most notably household income (question F41), father's education level (question F52), and mother's education level (question F56). In addition, the rotating module concerning democracy (questions E1-E45) appears to have contained some difficult questions. The income item nonresponse indicator, the parents' education item nonresponse indicator, and the democracy item nonresponse indicator equal 1 if an answer is missing ("don't know" or "refusal") for the related questions.

Multiple response is measured by a binary indicator that equals 1 if only one answer was recorded to an "all that apply" question. The question about activities conducted in the week prior to the interview (F17a) is the only non-conditional

question of this type in the ESS questionnaire. The respective show card mentions eight activities, such as paid work and education. The respondent could select all the activities that were applicable to him or her during the relevant week, and the interviewer was explicitly requested to probe for any other answers.

As an approximate indicator of differentiation, the amount of variability in the provided answers (i.e. the average standard deviation over all the items) is calculated for three sets of items covering three very different topics, namely the CES-D 8 depression scale (questions D5-D12), the understandings of democracy items (questions E1-E16), and the Schwarz Human Values scale (questions HF1/2).

Interview speed is calculated as the number of questions completed per minute (Loosveldt and Beullens 2013) in the main questionnaire. Although low interview speed in the first few interviews could be caused by a lack of familiarity with the survey instrument (Peytchev and Olson 2007), the interviewers involved in the ESS can be assumed to have been sufficiently familiar with the questionnaire, because they self-administered it at least once in the "task as respondent." Interview speed can therefore be interpreted as a sign of (in)accuracy, more than as a lack of familiarity with the questionnaire.

Modeling Approach.

In order to test the hypothesized associations between the response characteristics of interviewers and respondents, a two-level random intercept model, with respondents nested within interviewers, is estimated for each of the eight indicators. Each model has one of the respondent response indicators as its outcome variable, respondent age, gender, and education as control variables, an interviewer random effect and the

corresponding interviewer response indicator as an additional independent variable. For example, respondent nonresponse to the "income" item is explained by interviewer nonresponse to the same "income" item in the "task as respondent", after controlling for respondent age, gender, and education. A Bernoulli distribution is assumed for the four binary outcomes (the item nonresponse indicators and the multiple response indicator), whereas the four continuous outcomes (the differentiation indicators and the speed indicator) assume a normal distribution. The proportions of the variance in the response indicators explained by the interviewers (intraclass correlation, ICC) is estimated from the multilevel models, which include the respondent background characteristics–but before adding the interviewers' response behavior–to partially control for different respondent groups within interviewers (Hox 1994). For the models with a binary outcome, the residual variance component is fixed at $\frac{\pi^2}{3}$ (Goldstein, Browne, and Rasbash 2002).

Before reporting the estimated interviewer proportions of explained variance in the respondents' response characteristics and the estimated effects of the interviewers' response characteristics on the respective respondents' characteristics, we provide some descriptive statistics for the two groups. Even though the interviewers and respondents are not perfectly comparable due to the difference in administration mode, the respondent group provides a rough but useful benchmark, against which the interviewers' response characteristics can be evaluated.

Results

The descriptive statistics for the selected response indicators show mostly trivial differences between the interviewers and the respondents (Table 1), suggesting that

the interviewers put real effort into completing the "task as respondent". First, with the exception of the income question, the levels of item nonresponse in the interviewer group are low and comparable with those of the respondent group, indicating that the interviewers did not excessively select "don't know" or "refusal," even though these options were explicit in the self-administered mode. Second, the vast majority of the interviewers selected only one of the proposed options to the "all that apply" question, as did most of the respondents. Third, although the circumstances of the task–self-administration and no explicit request to answer questions truthfully–might have invited straight-lining behavior, the interviewers rarely produced pure straight lines (data not shown). On the contrary, they differentiated between answers to a similar extent as did the respondents. Lastly, the average interview speed for the interviewer group was higher than the speed for the respondent group, but not unreasonably so. Interview speed is expected to be somewhat higher for a self-administered questionnaire than for a personal interview. The pattern of speed over the different modules in the questionnaire is also comparable in both groups (data not shown).

Table 1: Descriptive Statistics of Response Indicators for Respondents and Interviewers

	Respondents			Interviewers		
	Mean	**SD**	**N**	**Mean**	**SD**	**N**
Item nonresponse, income	0.0877	0.2830	1869	0.6000	0.4915	155
Item nonresponse, parents' education	0.1113	0.3146	1869	0.0774	0.0215	155
Item nonresponse, democracy	0.1182	0.3230	1869	0.1419	0.0281	155
Multiple response	0.7956	0.4034	1869	0.7677	0.0340	155
Differentiation, depression scale	0.9674	0.2834	1868	0.8520	0.3470	154
Differentiation, values scale	1.1142	0.3235	1867	1.0580	0.0250	154
Differentiation, democracy	1.5756	0.7978	1862	1.6696	0.0504	154
Interview speed	3.8994	0.8323	1868	5.0761	1.9110	155

Note: SD = standard deviation; N = number of available cases.

The intra-class correlations (ICC), after controlling for the respondents' background characteristics (Table 2), confirm that the observed response characteristics for the respondents are to some extent due to the interviewers. Although the interviewers have only a modest effect on the variability of answers when questions are presented in a grid format (differentiation indicators), they strongly affect whether or not a substantive answer is recorded (item non response indicators), whether multiple answers are recorded, and the speed at which the interview is conducted.

Table 2: Interviewer Proportion of Explained Variance of Respondents' Suboptimal Response Behaviors

	ICC
Item nonresponse, income	0.177
Item nonresponse, parents' education	0.150
Item nonresponse, democracy	0.229
Multiple response	0.528
Differentiation, depression scale	0.050
Differentiation, values scale	0.061
Differentiation, democracy	0.027
Interview speed	0.273

Table 3 presents the estimated effect for each interviewer response indicator on the equivalent indicator for the respondents. These effects of the same interviewer characteristic are positive and statistically significant for income item nonresponse, multiple response to an "all that apply" question, and interview speed. The odds of a respondent having provided a response to the question about income doubles (exp(0.726) = 2.067) when the interviewer also gave an answer to this question. Similarly, the odds of a respondent providing multiple answers is almost five times greater (exp(1.561) = 4.7636) when the interviewer also selected multiple answers, and a respondent's interview speed is significantly higher when interviewed by an interviewer who answered questions faster. For the Schwarz' Human Values scale, the same-characteristic effect is also statistically significant, with interviewers who provided more variable answers in the "task as respondent" administering interviews

with more variable answers. There is no such effect for the two other differentiation indicators (depression scale and democracy module), or for the two other item nonresponse indicators (parents' education and democracy module).

Table 3: Estimated Effects and Standard Errors (SE) of Interviewer Response Characteristics–Included Separately after Controlling for Respondent Age, Gender, and Education level–on the Corresponding Respondent Response Characteristic

	Estimate	SE	p-value
Item nonresponse, income	0.726	0.253	0.004
Item nonresponse, parents' education	0.035	0.373	0.925
Item nonresponse, democracy	0.566	0.324	0.081
Multiple response	1.561	0.407	< 0.001
Differentiation, depression scale	0.037	0.025	0.138
Differentiation, values scale	0.100	0.032	0.001
Differentiation, democracy	0.003	0.035	0.941
Interview speed	0.091	0.020	< 0.001

Conclusion and Discussion

In this article, we examine the way in which interviewers answer surveys questions in a "task as respondent" and propose that interviewers who put little effort into this task will similarly expend little effort when interacting with respondents in the actual interview process. We find evidence for some of the hypothesized associations between interviewer and respondent response characteristics, specifically for income item nonresponse, multiple response, and interview speed. These findings suggest that first, interviewers who do not provide a substantive answer to the survey question regarding household income, subsequently administer more interviews in which a substantive answer to this question is also missing. Second, that interviewers who spend less time on the "task as respondent" administer faster interviews. Third, that interviewers who select only one of the answers in "all that apply" questions administer more interviews where only one answer is selected, and fourth, that interviewers who provide more variable answers to the Schwarz Human Values scale also administer interviews with more variability on this scale. The results for item nonresponse and response variability are mixed, with no statistically significant associations found for item nonresponse to questions concerning parents' education or the democracy module, and response variability in the CES-D 8 depression scale and in the democracy understanding list of items.

The estimated interviewer effects on item nonresponse are of considerable magnitude (in the range of 15-30%) and–in line with previous studies (Pickery and Loosveldt 2001)–support the claim that item nonresponse is not only attributable to the respondent and the questionnaire, but may also result from interviewers deviations during the interviewing process, such as erroneous skipping and inadequate probing

(de Leeuw, Hox, and Huisman 2003). In contrast to the results obtained by Pickery and Loosveldt (2001), we find that interviewers' own item nonresponse is indicative of respondents' item nonresponse only for the household income question–a particularly sensitive question–but not for the two other item nonresponse indicators. The effect of interviewers' item nonresponse to the income question on respondents' item nonresponse to the same question can be related to the study of Singer and Kohnke-Aguirre (1979), which reports consistently higher item nonresponse when interviewers anticipate difficulties. Interviewers who have little difficulty or restraint in reporting their own household income may be systematically better at asking the question in a neutral way without apologizing, and/or probing more intensively.

Inadequate probing could also explain the large interviewer effect on multiple response to "all that apply" questions. With regard to this question, interviewers are explicitly instructed to probe for more answers, allowing for interviewer differences in the extent of probing before continuing with the next question (Fowler and Mangione 1990). The effect of interviewers' own multiple response on respondents' multiple response suggests that interviewers who report only one activity probe very little or not at all, and that their respondents are therefore much more likely to select only one activity. The absence of multiple answers to "all that apply" questions therefore seems to capture interviewer satisficing more than it captures respondents' lack of motivation to answer the questions as completely as possible.

The estimated interviewer effect on interview speed (about 30%) is also large, consistent with the findings of Olson and Peytchev (2007) and Loosveldt and Beullens (2013). Further, the effect of interviewers' own interview speed on that of respondents is strong and statistically significant. This association can be explained by

some interviewers wanting to finish their tasks faster, for instance by reading questions more quickly, which will be reflected in both the "task as respondent" and in the actual survey interviews administered.

The evidence concerning the association between interviewer and respondent response variability is mixed, with only one set of items for which a statistically significant association is found. It is unclear why there would only be such an association for this list of items. One way in which this differs from the other two is that it appears at the end of the interview. A possible explanation for the association could therefore be that some interviewers consider this last module of secondary importance and spend less effort when asking these questions. Nevertheless, interviewer effects on response variability for this list of items, as well as the other list of items studied, are found to be relatively small (explained variance in the range of 3-6%).

As illustrated by this study, the "task as respondent"–in which interviewers self-administer the questionnaire–is not only valuable for interviewers to become familiar with the survey instrument, but also provides useful information about the interviewers at almost no additional cost. The substantive data from the "task as respondent" does not necessarily reflect the interviewers' socio-demographics, and their socio-political opinions and attitudes, because the interviewers are not explicitly requested to answer the questions truthfully, and concerns about privacy may arise if interviewers feel forced to provide (sensitive) information that they would otherwise not divulge. However, formal measurements of response behavior (e.g. item nonresponse, multiple response, differentiation, and interview speed) can be derived,

and this information about the manner in which interviewers answer survey questions in the "task as respondent" can be taken advantage of.

Our results show that it is possible to predict, to some extent, that the survey interviews administered by a given interviewer might exhibit some specific response characteristics, based solely on the data from the "task as respondent" before the relevant interviewer administered any interviews. These results suggest that characteristics of respondents' answers, which are usually attributed to the satisficing behavior of respondents, are to some extent due to the satisficing behavior of interviewers.

Because these associations are likely to be due to the manifestation of particular interviewer deviations during the survey interviews, the results of the "task as respondent" may be used with regard to feedback and training. The interviewers may be given–before they start their interviewing task–feedback that is to some extent specific to their respective likely deviations. For example, the importance of speaking slightly slower than normal conversation speed may be stressed to interviewers who carry out the "task as respondent" unusually fast. The importance of probing in the case of multiple-choice multiple-answer questions may be stressed to interviewers who select only one of the answers. The fact that most respondents, when asked sensitive questions by an interviewer in a normal way without apologies do not mind providing an answer, may be stressed to interviewers who do not provide a meaningful answer themselves. Of course, all of these basic instructions are usually given to all the interviewers working on a particular project. Nevertheless, based on the "task as respondent," targeted individual positive or negative feedback for each interviewer could be useful in order to focus their attention on the instructions they

are most likely to diverge from. Similarly, project briefings and training may be adapted to focus on likely deviations. A third approach to dealing with interviewer performance, selection, may be considered, but the effect sizes estimated in our study do not seem to warrant removal of interviewers from the project solely on the basis of this task. In addition, selection would reduce the size of the interviewer workforce, and could also induce interviewers to do well on tasks they know are monitored while taking less care over tasks that remain unobserved. In order to limit (unintentional) interviewer deviations and improve overall performance, feedback and training are preferable. Further research is necessary to investigate the effectiveness of such targeted feedback and training programs.

References

Ackermann-Piek, D., Massing, N. 2014. Interviewer Behavior and Interviewer Characteristics in PIAAC Germany. *Methods, Data, Analyses* 8(2): 199–222. doi:10.12758/mda.2014.008.

Billiet, J., Loosveldt, G. 1988. Improvement of the Quality of Responses to Factual Survey Questions by Interviewer Training. *The Public Opinion Quarterly* 52(2):190–211.

De Leeuw, E. D., Hox, J., Huisman, M. 2003. Prevention and Treatment of Item Nonresponse. *Journal of Official Statistics* 19(2):153–76.

Dijkstra, W., Ongena, Y. 2006. Question-Answer Sequences in Survey-Interviews. *Quality & Quantity* 40(6):983–1011. doi:10.1007/s11135-005-5076-4.

European Social Survey. 2012. *ESS Round 6 Data File Edition 2.1*. Norway: Norwegian Social Science Data Services, Data Archive and distributor of ESS data.

Cannell, C. F., Fowler, F. J., Marquis, K. H. 1968. The Influence of Interviewer and Respondent Psychological and Behavioral Variables on the Reporting in Household Interviews. *Vital and Health Statistics* 2(26).

Fowler, F. J., Mangione, T. W. 1990. *Standardized Survey Interviewing*. California: SAGE.

Goldstein, H., Browne, W., Rasbash, J. 2002. Partitioning Variation in Multilevel Models. *Understanding Statistics* 1:223–231.

Hox, J. J. 1994. Hierarchical Regression Models for Interviewer and Respondent Effects. *Sociological Methods & Research* 22(3):300–318. doi:10.1177/0049124194022003002.

Houtkoop-Steenstra, H. 2000. *Interaction and the Standardized Survey Interview: The Living Questionnaire*. Cambridge/New York: Cambridge University Press.

Japec, L., ed. 2008. Interviewer Error and Interviewer Burden. Pp. 187–211 in *Advances in Telephone Survey Methodology*, edited by J. M. Lepkowski, C. Tucker, J. M. Brick, E. D. de Leeuw, L. Japec, P. J. Lavrakas, M. W. Link, R. L. Sangster. Hoboken, NJ: John Wiley & Sons, Wiley Series in Survey Methodology.

Krosnick, J. 2000. The Threat of Satisficing in Surveys: The Shortcuts Respondents Take in Answering Questions. *Survey Methods Newsletter* 20(1).

Krosnick, J. A., Narayan, S., Smith, W. R. 1996. Satisficing in Surveys: Initial Evidence. *New Directions for Evaluation* (70): 29–44. doi:10.1002/ev.1033.

Krosnick, J. A., Alwin, D. F. 1987. Satisficing: A Strategy for Dealing with the Demands of Survey Questions. *GSS Methodological Report* 46.

Loosveldt, G., Beullens, K. 2013. The Impact of Respondents and Interviewers on Interview Speed in Face-to-Face Interviews. *Social Science Research* 42(6):1422–30. doi:10.1016/j.ssresearch.2013.06.005.

Mangione, T. W., Fowler, F. J., Louis, T. A. 1992. Question Characteristics and Interviewer Effects. *Journal of Official Statistics* 8 (3):293–307.

O'Muircheartaigh, C., Campanelli, P. 1998. The Relative Impact of Interviewer Effects and Sample Design Effects on Survey Precision. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 161 (1): 63–77. doi:10.1111/1467-985X.00090.

Olson, K., Peytchev, A. 2007. Effect of Interviewer Experience on Interview Pace and Interviewer Attitudes. *Public Opinion Quarterly* 71(2):273–86. doi:10.1093/poq/nfm007.

Ongena, Y. P., Dijkstra, W. 2007. A Model of Cognitive Processes and Conversational Principles in Survey Interview Interaction. *Applied Cognitive Psychology* 21 (2): 145–63. doi:10.1002/acp.1334.

Pickery, J., Loosveldt, G. 1998. The Impact of Respondent and Interviewer Characteristics on the Number of 'No Opinion' Answers: A Multilevel Model for Count Data. *Quality & Quantity* 32:31–45.

Pickery, J., Loosveldt, G. 2001. An Exploration of Question Characteristics That Mediate Interviewer Effects on Item Nonresponse. *Journal of Official Statistics* 17(3):337–50.

Rasinski, K. A., Mingay, D., Bradburn, N. M. 1994. Do Respondents Really 'Mark All That Apply' on Self-Administered Questions? *Public Opinion Quarterly* 58:400–408.

Simon, H. A. 1957. *Models of Man: Social and Rational-Mathematical Essays on Rational Human Behavior in a Social Setting*. New York: Wiley.

Schober, M. F., Conrad, F. G. 1997. Does Conversational Interviewing Reduce Survey Measurement Error? *Public Opinion Quarterly* 61 (4): 576–602. doi:10.1086/297818.

Singer, E., Kohnke-Aguirre, L. 1979. Interviewer Expectation Effects: A Replication and Extension. *Public Opinion Quarterly* 43 (2): 245–60.

Tourangeau, R., Rips, L. J., Rasinski, K. A. 2000. *The Psychology of Survey Response*. New York: Cambridge University Press. doi:10.1017/CBO9780511819322.

West, B. T., Kreuter, F., Jaenichen, U. 2013. 'Interviewer' Effects in Face-to-Face Surveys: A Function of Sampling, Measurement Error, or Nonresponse? *Journal of Official Statistics* 29(2). doi:10.2478/jos-2013-0023.